9p.

003 Goddard

N64-20685

CODE-1

NASA CR-56098

CAT. 20

SYSTEMS ANALYSIS AND EVALUATION
TECHNICAL NOTE NO. 9
THE METHOD OF ADJOINTS AS APPLIED TO DETERMINISTIC LINEARIZED SYSTEMS

11 December 1963

MARTIN RUSH

CONTRACT NO. NAS-5-3486

INTERNATIONAL BUSINESS MACHINE CORPORATION
SPACE SYSTEMS DEPARTMENT
BETHESDA, MARYLAND

R64-15383

ABSTRACT

TITLE: The Method of Adjoints as Applied to Deterministic Linearized Systems

DATE: 11 December 1963

AUTHOR: Martin Rush

20685 A

This note presents an analytical development of the adjoint method as applied to linear deterministic systems. The development utilizes matrix notation as much as possible to simplify the development. The note also includes a simple example of the method to outline the computational steps required.

Author

1. Introduction

The method of adjoints is a powerful analytical tool in the determination of the sensitivities of a system's final state to perturbations in the control or state of the system along its trajectory in the state space. This technical report provides an analytical development of the adjoint method. An example of the method's application to deterministic linear systems is used to outline the procedure.

2. Linear Systems

We shall study the linear dynamical system

$$\frac{dx}{dt} = F(t)x + G(t)\,u(t) \qquad x(t_o) = C \tag{1.1}$$

where:

t = time
x = an n vector, the system state
$u(t)$ = an m vector, the control
$F(t)$ = an nxn matrix
$G(t)$ = an nxm matrix

Equation (1.1) may represent a model of a linear system or a linearized model of a nonlinear system. All quantities in (1.1) will be real. The set of all x is thus a real $\underline{n}$-dimensional Euclidean space called the state space and denoted by $\underline{X}$. If $F(t)$ & $G(t)$ are constant the system is called constant or time invariant. If $u(t) = 0$ the system is called free and we have a set of homogeneous equations. The functions $F(t)$, $G(t)$, $u(t)$ are defined for all $-\infty < t < \infty$ and are bounded in every bounded interval of t. In addition we assume throughout that they are measurable functions of t.

Let us form the "fundamental solution matrix."

$$x = \Phi(t, t_o)\, y$$

where

$$\frac{d}{dt}\Phi(t, t_o) = F(t)\,\Phi(t, t_o) \qquad \text{with } \Phi(t_o, t_o) = I \,. \tag{1.2}$$

$\Phi(t, t_o)$ is the matrix formed by n independent solutions of the homogeneous part of (1.1), each solution being obtained by a unit initial condition in the mth component of the x vector and setting the other (n-1) components of x to zero at t_o. Then n solutions are arranged in a matrix.

$$\Phi(t, t_o) = \begin{bmatrix} x_1^{(1)}(t) \ldots\ldots x_1^{(m)}(t) \ldots\ldots x_1^{(n)}(t) \\ x_2^{(1)}(t) \ldots\ldots x_2^{(m)}(t) \ldots\ldots \\ x_n^{(1)}(t) \ldots\ldots x_n^{(m)}(t) \ldots\ldots x_n^{(n)}(t) \end{bmatrix} \quad (1.3)$$

where the superscript denotes the solution and the subscript the component of the vector. Furthermore the initial conditions are

$$\Phi(t_o, t_o) = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & & \vdots & & \vdots \\ & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cdots & 1 & \cdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} = I \quad (1.4)$$

as given in (1.2).

Substituting (1.2) in (1.1):

$$\frac{d\,\Phi(t, t_o)}{dt}\, y(t) + \Phi(t, t_o)\frac{dy(t)}{dt} = F(t)\,\Phi(t, t_o)\, y + G(t)\, u(t) \quad (1.5)$$

hence:

$$\frac{dy(t)}{dt} = \Phi^{-1}(t, t_o)\, G(t)\, u(t) \quad (1.6)$$

$$y(t) = C + \int_{t_o}^{t} \Phi^{-1}(\tau, t_o)\, G(\tau)\, u(\tau)\, d\tau \,. \quad (1.7)$$

Substituting in (1.2)

$$x(t) = \Phi(t, t_o)\, C + \int_{t_o}^{t} \Phi(t, t_o)\, \Phi^{-1}(\tau, t_o)\, G(\tau)\, u(\tau)\, d\tau \,. \quad (1.8)$$

Now:
$$\Phi(t, t_o) = \Phi(t, \tau)\,\Phi(\tau, t_o)$$

$$\therefore \quad x(t) = \Phi(t, t_o)\, C + \int_{t_o}^{t} \Phi(t, \tau)\, G(\tau)\, u(\tau)\, d\tau \, . \qquad (1.9)$$

Since $x(t_o) = y(t_o)$ from (1.2) we may substitute $x(t_o)$ for C in equation (1.9) and obtain:

$$x(t) = \Phi(t, t_o)\, x(t_o) + \int_{t_o}^{t} \Phi(t, \tau)\, G(\tau)\, u(\tau)\, d\tau \qquad (1.10)$$

where:

$$\frac{dx}{dt} = F(t)\, x + G(t)\, u(t) \qquad (1.11)$$

Equation (1.10) may be used to obtain the effects of initial conditions of a set of linear or linearized equations on the state vector at some t . The first term in (1.10) provides the means of obtaining the sensitivities of the state vector x(t) to variations in the initial state $x(t_o)$. The second term provides the means of determining the effects of the control u on the final state x(t). For an impulsive perturbation in the control at some time τ we would then have:

$$\delta x(t) = \Phi(t, \tau)\, G(\tau)\, \delta u \qquad (1.12)$$

$$\text{for } t \geq \tau$$

where $\delta x(t)$ is the perturbation in the final state and δu is the impulsive perturbation in the control at time τ . It becomes evident; however, that the utilization of equation set (1.10) may be a cumbersome job. To determine the effects of disturbing forces along a descent trajectory during re-entry would require many solutions of equation 1.2 for various t_o's.

$$\frac{d\Phi(t, t_o)}{dt} = F(t)\, \Phi(t, t_o) \qquad \Phi(t_o, t_o) = I \qquad (1.13)$$

to obtain:

$$\frac{\delta x(t)}{\delta x(t_o)} = \Phi(t, t_o) \, . \qquad (1.14)$$

Similarly the effect of the control as applied at $t = \Upsilon$ requires the determination of $\Phi\,(t, \Upsilon)$ for $t_0 \leq \Upsilon \leq t$, again a large simulation job. There is happily a more efficient method of attacking this problem by means of adjoint functions.

3. The Adjoint Method

Let us attempt to find a solution matrix adjoint to 1.13:

$$\Phi(\Upsilon, t)\,\Phi^{-1}(\Upsilon, t) = I \tag{2.1}$$

where $\Phi\,(t, t) = I$.

Differentiating (2.1) with respect to Υ.

$$\Phi\,(\Upsilon, t)\frac{d}{d\Upsilon}\Phi^{-1}(\Upsilon, t) + \left[\frac{d}{d\Upsilon}\Phi\,(\Upsilon, t)\right]\Phi^{-1}(\Upsilon, t) = 0\,. \tag{2.2}$$

Utilizing (1.13)

$$\frac{d}{d\Upsilon}\Phi\,(\Upsilon, t) = F(\Upsilon)\,\Phi\,(\Upsilon, t)\,. \tag{2.3}$$

Substituting (2.3) in (2.2).

$$\frac{d}{d\Upsilon}\Phi^{-1}(\Upsilon, t) = -\Phi^{-1}(\Upsilon, t)\,F(\Upsilon) \tag{2.4}$$

since:

$$\Phi^{-1}(\Upsilon, t) = \Phi\,(t, \Upsilon) \tag{2.5}$$

$$\frac{d}{d\Upsilon}\Phi\,(t, \Upsilon) = -\Phi\,(t, \Upsilon)\,F(\Upsilon) \tag{2.6}$$

or transposing:

$$\frac{d}{d\Upsilon}\Phi^T(t, \Upsilon) = -F^T(\Upsilon)\,\Phi^T(t, \Upsilon) \tag{2.7}$$

with $\Phi\,(t, t) = I$.

Equation 2.7 is the solution matrix of a set of differential equations adjoint to (1.2).

Let

$$\lambda(\Upsilon) = \Phi^T(t, \Upsilon)\,\lambda\,(t) \tag{2.8}$$

then $\lambda(\Upsilon)$ satisfies

$$\frac{d}{d\Upsilon}\lambda\,(\Upsilon) = -F^T(\Upsilon)\lambda(\Upsilon)\,. \tag{2.9}$$

$\lambda(\Upsilon)$ represents an n vector adjoint to $x(t)$ and $\Phi^T(t,\Upsilon)$ a solution matrix adjoint to $\Phi\,(\Upsilon, t_0)$.

Thus by solving equation set (2.7) only once we can obtain Φ (t, Υ) required in equation set (1.10).

4. Example of the Adjoint Method

As an example consider a particle falling vertically under a constant gravitation acceleration g:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -g \end{aligned} \tag{3.1}$$

where x_1 is distance and x_2 velocity.
Referring to equation set (1.11).

$$F(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad G(t) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.2}$$

and

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -g \end{bmatrix}. \tag{3.3}$$

By normal integration from t_0 to t

$$\begin{aligned} x_1(t) &= x_1(t_o) + x_2(t_o)\,(t-t_o) - \frac{g}{2}\,(t-t_o)^2 \\ x_2(t) &= x_2(t_o) - g(t-t_o) \end{aligned} \tag{3.4}$$

and taking the partials of (3.4).

$$\frac{\partial x_1(t)}{\partial x_1(t_o)} = 1 \qquad \frac{\partial x_1(t)}{\partial x_2(t_o)} = t-t_o \tag{3.5}$$

$$\frac{\partial x_2(t)}{\partial x_1(t_o)} = 0 \qquad \frac{\partial x_2(t)}{\partial x_2(t_o)} = I.$$

Using the adjoint technique:

$$\dot{\Phi}^T(t,\tau) = -F^T(\tau)\,\Phi^T(t,\tau) \tag{2.7}$$

$$\Phi^T(t,t) = I$$

or

$$\begin{bmatrix} \dot{\Phi}_{11} & \dot{\Phi}_{21} \\ \dot{\Phi}_{12} & \dot{\Phi}_{22} \end{bmatrix} = -\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{21} \\ \Phi_{12} & \Phi_{22} \end{bmatrix} \tag{3.6}$$

with $\Phi(t,t) = I$

then:

$$\begin{aligned} \dot{\Phi}_{11} &= 0 & \Phi_{11}(t,\tau) &= \Phi_{11}(t,t) \\ \dot{\Phi}_{12} &= -\Phi_{11} & \Phi_{12}(t,\tau) &= \int_t^{\tau} -\Phi_{11}(t,t)\,d\tau + \Phi_{12}(t,t) \\ \dot{\Phi}_{21} &= 0 & \Phi_{21}(t,\tau) &= \Phi_{21}(t,t) \\ \dot{\Phi}_{22} &= -\Phi_{21} & \Phi_{22}(t,\tau) &= \int_t^{\tau} -\Phi_{21}(t,t)\,d\tau + \Phi_{22}(t,t) \end{aligned} \tag{3.7}$$

and from the definition of $\Phi(t,t)$:

$$\begin{aligned} \Phi_{11}(t,t) &= 1 \\ \Phi_{12}(t,t) &= 0 \\ \Phi_{21}(t,t) &= 0 \\ \Phi_{22}(t,t) &= 1 \end{aligned} \tag{3.8}$$

or:

$$\Phi(t,\tau) = \begin{bmatrix} 1 & t-\tau \\ 0 & 1 \end{bmatrix} \tag{3.9}$$

and from (1.10)

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & t-t_o \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t_o) \\ x_2(t_o) \end{bmatrix} + \int_{t_o}^{t} \begin{bmatrix} 1 & t-\tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -g \end{bmatrix} d\tau \tag{3.10}$$

It is observed that equation (3.10) provides the sensitivities of $x(t)$ to perturbations in $x(t_o)$, giving the same results as (3.5) equation set (3.5)

REFERENCES:

1. Tsien, H.S., "Engineering Cybernetics", McGraw-Hill Book Company Inc. N.Y. 1954

2. Bellman, R., "Introduction to Matrix Analysis", McGraw-Hill Book Company, Inc. N.Y. 1960